

10

Presència del
català a la xarxa

10.1. Presència del català a Internet

Indicadors Recursos i serveis lingüístics en català a la xarxa.
Cercadors que permeten la cerca en català.
Pàgines i llocs web en català.

Conceptes Es consideren eines lingüístiques en català tota mena de programari que serveix per utilitzar paquets informàtics fent servir aquesta llengua, o bé que permeti transformar informació escrita o oral d'una llengua determinada al català o a la inversa. Això darrer inclou eines de traducció automàtica i de síntesi i generació de llenguatge oral.

Pel que fa a la presència del català a la xarxa, és essencial saber quants organismes catalans (empreses i entitats) són presents a la xarxa i quina és la rellevància dels seus continguts en línia. Com a primera aportació per a futurs estudis que es puguin realitzar sobre els continguts en llengua catalana presents a la Web (no a intranets ni a extranets) s'ha fet una cerca de tots els llocs i pàgines web en català. Per fer aquesta cerca s'ha utilitzat un robot lingüístic seguint les següents pautes:

- Selecció de lèxic català adequat que permet identificar pàgines web com a pertanyents a l'àmbit lingüístic català (valencià i balear inclosos).
- Programació dels robots per al rastreig de la web, emprant tecnologies de Metacrawling (contra el cercador Google), i per a l'anàlisi de lligams creuats a partir de *links* llavors que han permès arribar a *links* referits de forma recurrent i filtrant informació per tal d'obtenir només pàgines en català.
- Enregistrament de les pàgines web, i els seus DNS (URL).
- Els *links* de *links* obtinguts, si apuntaven a adreces ja donades d'alta a les bases de dades, incrementaven un comptador de reincidència, cosa que permet calcular quins *links* han estat més referits des d'altres; per tant, s'usa el mètode del *link* més apuntat com a nivell de la bondat, qualitat o "èxit" del *link* referit.

Aquest estudi ens ha permès, finalment, disposar del nombre de recursos web en català (pàgines HTM, SHTML, PHP, ASP, PDF, JavaScript, JAVA,...). S'ha de considerar, però, que aquesta xifra no representa tota la informació disponible a Internet en català ja que molts continguts de la Web només són accessibles via formularis, o bé requereixen permís per accedir-hi (nom d'usuari i mot de pas), o que són consultables via llenguatges de programació interpretats, pàgines dinàmiques, etc.

Dada clau **El servei intern de la Generalitat de traducció automàtica català-castellà i viceversa ha traduït, en 4 mesos, prop de 7.700 pàgines (2,3 milions de paraules). A Internet hi ha 1.682.600 pàgines web en català dins de 47.241 llocs web.**

Informació de base *Recursos*

La següent llista presenta alguns dels llocs Internet més rellevants en què es troben recursos en llengua catalana:

- cultura.gencat.es/llengcat/informat Departament de Cultura
- www.cpnl.org/recursos Consorci per a la Normalització Lingüística

- www.softcatala.org Softcatalà
- www.llelengcat.com Directori virtual de la llengua catalana
- www.termcat.es Centre de terminologia en català
- www.iec.es Institut d'Estudis Catalans
- www.grec.net Hiperenciclopèdia. Diccionari de la llengua catalana
- sofia.uji.es/software Xarxa d'Universitats Institut Joan Lluís Vives

A l'adreça del Departament de Cultura es pot trobar informació sobre el programari en català, fabricants, material informàtic per aprendre català, etc. La web del Consorci per a la Normalització Lingüística permet obtenir assessorament lingüístic en línia, diccionaris, models de documents i altres recursos. A www.llelengcat.org es poden trobar portals en català, cercadors, jocs, campanyes, revistes, informàtica, serveis lingüístics, entitats que treballen per la llengua, etc.

Altres llocs web en català que hi ha a la Xarxa són: www.vilaweb.com, www.racocatala.com, www.andorra.ad, www.quands.com, www.estelnet.com, www.elbalco.com, www.girona.com i www.guiauep.com.

Traductors automàtics

Al mercat hi ha traductors automàtics amb el català com a llengua d'origen o destinació. Els serveis de traducció en línia en català es troben a:

- www.incyta.es
- www.automatictrans.es
- www.internostrum.com

La Generalitat, a través de la Secretaria de Telecomunicacions i Societat de la Informació ha posat en marxa el servei de traducció automàtica de documents per mitjà de la seva Intranet. D'aquesta manera tots els departaments i organismes que en depenen poden traduir documents entre el castellà i el català¹. El sistema disposa d'un servidor central que realitza i distribueix les traduccions quasi en temps real i amb una qualitat molt alta.

El servei es va posar en marxa el març del 2001 i fins al juny s'havien traduït en línia 7.700 pàgines (més de 2,3 milions de paraules). Es pot considerar, però, que actualment el servei està encara en la seva fase de desplegament, cosa que, per a molts departaments, significa utilitzar el servei en "fase de proves" i no a ple rendiment.

Cercadors

Molts cercadors permeten, des de fa temps, fer cerques en diferents idiomes, entre ells el català (Terra, All The Web, Google, etc.). De fet, el nombre de cercadors que permeten la cerca de termes i conceptes en llengua catalana s'ha incrementat molt darrerament. A continuació hi ha una llista amb diferents cercadors en català. Alguns d'aquests són exclusivament en català, com per exemple Cercat, Calalunya *on-line*, Cercador.com, etc.

¹ El servei intern de traducció de la Generalitat espera disposar del traductor automàtic entre el català i l'anglès a la tardor del 2001.

• All The Web	http://www.alltheweb.com
• Areabase	http://www.areabase.com
• Catalunya <i>on line</i>	http://www.catalunyaonline.com/
• Cercador.com	http://www.grec.net/cgi-bin/cercador.pgm
• CerCat	http://www.cercat.com/
• Dredera.com	http://www.dredera.com/
• Eureka	http://www.eureka.creativeweb.es/
• Google	http://www.google.com/
• Indibil - Webs de Ponent	http://webs.lleida.net/
• Infolleida - La guia de Lleida	http://www.infolleida.com/
• La Tecla	http://www.latecla.com/
• La Teranyina Gironina	http://www.teranyina.net/
• MallorcaWeb	http://www.mallorcaweb.com/
• Nosaltres.com	http://vilaweb.com/nosaltres/
• PaisosCatalans.com	http://www.paisoscatalans.org/
• Quèstia: cercador andorrà	http://monvirtual.com/questia/
• Sabadellencs	http://www.sabadellencs.com/
• Som-hi	http://www.som-hi.com/
• Terra	http://www.terra.es/
• Uep!	http://www.guiaupe.com/

Nombre de pàgines indexades en català

D'acord amb un estudi realitzat per la Secretaria de Telecomunicacions i Societat de la Informació,² el nombre de pàgines en català a Internet era de 1.682.600 a maig de 2001, i el nombre d'arrels de dominis que tenen algun contingut en català era de 47.241 llocs web.

Seguint els mateixos criteris s'han trobat per al castellà 23.000.000 pàgines per les 1.682.579 en català ja esmentades. Segons els cercador Google, a juny de 2001 hi havia unes 1.346.966.000 pàgines web en qualsevol idioma.

Les webs amb més contingut, les *top 15*, són:

- | | |
|---|---|
| • www.xtec.es (1.974 pàgines) | • www.uv.es (409 pàgines) |
| • www.avui.com (1.742 pàgines) | • www.pangea.org (395 pàgines) |
| • www.gencat.es (1.449 pàgines) | • www.verba.org (352 pàgines) |
| • www.fut.es (1.038 pàgines) | • www.partal.com (345 pàgines) |
| • www.ub.es (689 pàgines) | • www.sic.uji.es (333 pàgines) |
| • www.geocities.com (597 pàgines) | • www.tvcatalunya.com (319 pàgines) |
| • www.uib.es (591 pàgines) | • www.bcn.es (316 pàgines) |
| • www.wordtheque.com (472 pàgines) | |

² Aquest estudi ha estat encarregat a l'empresa ADDA, especialitzada en desenvolupament d'aplicatius i serveis lingüístics.

Per altra banda, el nombre de pàgines indexades que es poden trobar en català és, si es fa una cerca sense cap altre criteri que el de "llengua catalana" al cercador *All The Web* (www.alltheweb.com), a juliol 2001, de 757.911, mentre que a juny de 2000 era de 443.301 pàgines. L'increment de les pàgines en català respecte a l'any anterior és de més del 70%.

Pel que fa al castellà, al cercador *All The Web* es van trobar 7.573.064 pàgines a juny de l'any 2000, mentre que a juliol del 2001 ja eren 16.172.037. I en anglès n'hi havia 214.250.996 l'any 2000, i 332.771.089 el 2001.

10.2. Dominis Internet d'organismes i empreses amb seu a Catalunya

Indicador Nombre de dominis de primer nivell registrats per l'ES-NIC¹ a Catalunya.

Dada clau **El nombre de dominis ".es" d'empreses i organismes amb seu social a Catalunya va augmentar més d'un 50% entre 1999 i 2000.**

Informació de base A mitjan de 2001 hi havia 9.986 dominis ".es" registrats per l'ES-NIC a Catalunya². Aquesta xifra s'ha incrementat més d'un 38% respecte a l'any 2000, quan hi havia 7.236 dominis.

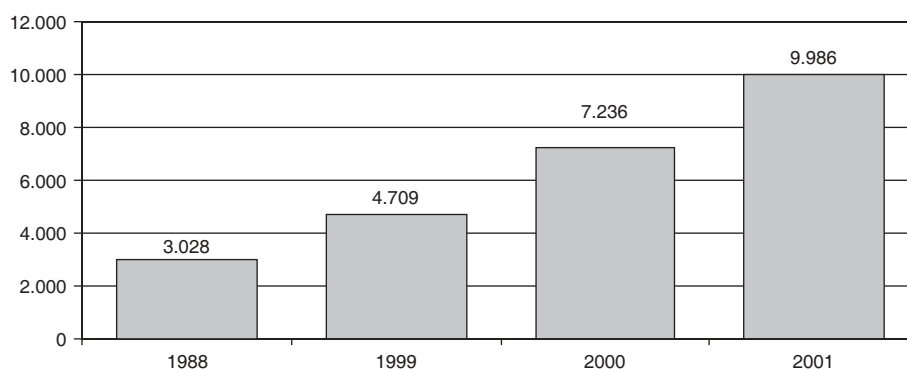
Els dominis presenten una clara concentració geogràfica, atès que a Barcelona hi tenen la seu més del 90% de les empreses amb domini ".es" de Catalunya. En conjunt, Catalunya té el 27,6% del total de dominis ".es" donats d'alta a tot Espanya (36.125).

DOMINIS ".ES" INSCRITS A CATALUNYA				
	Octubre 98	Octubre 99	Agost 00	Setembre 01
Barcelona	2.927	4.280	6.668	9.051
Girona	0	188	291	416
Lleida	0	83	27	187
Tarragona	101	158	250	332
Total Catalunya	3.028	4.709	7.236	9.986
% Cat. / Esp.	27,4%	28,7%	28,0%	27,6%
Total Espanya	11.055	16.434	25.844	36.125

Nota: Evolució de la inscripció de dominis sota ".es" de titulars amb domicili social a la Comunitat Autònoma de Catalunya.

Font: CMT, octubre 98, octubre 99, juny 00 i setembre 01.

DOMINIS ".ES" INSCRITS A CATALUNYA



Font: CMT, octubre 98, octubre 99 i juny 00, setembre 01.

¹ ES-NIC, Registre Delegat d'Internet a Espanya.

² Comisión del Mercado de las Telecomunicaciones.